

#7

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/668,482

DATE: 08/30/2001

TIME: 15:04:31

Input Set : N:\Crf3\RULE60\09668482.txt

Output Set: N:\CRF3\08302001\I668482.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Petkovich, P. Martin, White, Jay A.,
6 Beckett, Barbara R., Jones, Glenville
8 (ii) TITLE OF INVENTION: Retinoid Metabolizing Protein

10 (iii) NUMBER OF SEQUENCES: 43

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Blake, Cassels & Graydon

14 (B) STREET: Box 25, Commerce Court West

15 (C) CITY: Toronto

16 (D) STATE: Ontario

17 (E) COUNTRY: Canada

18 (F) ZIP: M5L 1A9

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage

22 (B) COMPUTER: COMPAQ, IBM PC compatible

23 (C) OPERATING SYSTEM: MS-DOS 5.1

24 (D) SOFTWARE: WORD PERFECT

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/09/668,482

C--> 28 (B) FILING DATE: 25-Sep-2000

30 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: 08/882,164

34 (B) FILING DATE: June 25, 1997

35 (A) APPLICATION NUMBER: 08/667,546

36 (B) FILING DATE: June 21, 1996

37 (A) APPLICATION NUMBER: 08/724,466

38 (B) FILING DATE: October 1, 1996

40 (viii) ATTORNEY/AGENT INFORMATION:

41 (A) NAME: Hunt, John C.

42 (B) REGISTRATION NUMBER: 36,424

43 (C) REFERENCE/DOCKET NUMBER: 50767/00010

45 (ix) TELECOMMUNICATION INFORMATION:

46 (A) TELEPHONE: (416) 863-4344

47 (B) TELEFAX: (416) 863-2653

50 (2) INFORMATION FOR SEQ ID NO: 1

51 (i) SEQUENCE CHARACTERISTICS:

52 (A) LENGTH: 337 base pairs

53 (B) TYPE: nucleic acid

54 (C) STRANDEDNESS: single

55 (D) TOPOLOGY: linear

57 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1

59	TGCCAGTGGA CAATCTCCCT ACCAAATTCA CTAGTTATGT CCAGAAATTA GCCTAAACCG	60
61	GAGCCTTTGT ACATATGTTT TTATTTTAGA TGAAGTGTGA TGTATTGGAT ATTTTCTAAT	120
63	TTGTTTATAT AAAGCAGATG TGTATATAAG TCTATGCGAA GAAGCGAAAA CGAGGGCACT	180
65	ACTTTCTCAT GGATCACTGT AATGCTACAG AGTGTCTGTG ATGTATATTT ATAATGTAGT	240
67	TGTGTCATAT AGCTTTTGTA CTGTATGCAA CTTATTTAAC TCGCTCTTTA TCTCATGGGT	300

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/668,482

DATE: 08/30/2001

TIME: 15:04:31

Input Set : N:\CrF3\RULE60\09668482.txt

Output Set: N:\CRF3\08302001\I668482.raw

69 TTTATTTAAT AAAACATGTT CTTACAAAAA AAAAAAA 337

72 (2) INFORMATION FOR SEQ ID NO: 2

73 (i) SEQUENCE CHARACTERISTICS:

74 (A) LENGTH: 492 amino acids

75 (B) TYPE: amino acid

76 (C) STRANDEDNESS: single

77 (D) TOPOLOGY: linear

79 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2

81 Met Gly Leu Tyr Thr Leu Met Val Thr Phe Leu Cys Thr Ile Val Leu

82 1 5 10 15

84 Pro Val Leu Leu Phe Leu Ala Ala Val Lys Leu Trp Glu Met Leu Met

85 20 25 30

87 Ile Arg Arg Val Asp Pro Asn Cys Arg Ser Pro Leu Pro Pro Gly Thr

88 35 40 45

90 Met Gly Leu Pro Phe Ile Gly Glu Thr Leu Gln Leu Ile Leu Gln Arg

91 50 55 60

93 Arg Lys Phe Leu Arg Met Lys Arg Gln Lys Tyr Gly Cys Ile Tyr Lys

94 65 70 75 80

96 Thr His Leu Phe Gly Asn Pro Thr Val Arg Val Met Gly Ala Asp Asn

97 85 90 95

99 Val Arg Gln Ile Leu Leu Gly Glu His Lys Leu Val Ser Val Gln Trp

100 100 105 110

102 Pro Ala Ser Val Arg Thr Ile Leu Gly Ser Asp Thr Leu Ser Asn Val

103 115 120 125

105 His Gly Val Gln His Lys Asn Lys Lys Lys Ala Ile Met Arg Ala Phe

106 130 135 140

108 Ser Arg Asp Ala Leu Glu His Tyr Ile Pro Val Ile Gln Gln Glu Val

109 145 150 155 160

111 Lys Ser Ala Ile Gln Glu Trp Leu Gln Lys Asp Ser Cys Val Leu Val

112 165 170 175

114 Tyr Pro Glu Met Lys Lys Leu Met Phe Arg Ile Ala Met Arg Ile Leu

115 180 185 190

117 Leu Gly Phe Glu Pro Glu Gln Ile Lys Thr Asp Glu Gln Glu Leu Val

118 195 200 205

120 Glu Ala Phe Glu Glu Met Ile Lys Asn Leu Phe Ser Leu Pro Ile Asp

121 210 215 220

123 Val Pro Phe Ser Gly Leu Tyr Arg Gly Leu Arg Ala Arg Asn Phe Ile

124 225 230 235 240

126 His Ser Lys Ile Glu Glu Asn Ile Arg Lys Lys Ile Gln Asp Asp Asp

127 245 250 255

129 Asn Glu Asn Glu Gln Lys Tyr Lys Asp Ala Leu Gln Leu Leu Ile Glu

130 260 265 270

132 Asn Ser Arg Arg Ser Asp Glu Pro Phe Ser Leu Gln Ala Met Lys Glu

133 275 280 285

135 Ala Ala Thr Glu Leu Leu Phe Gly Gly His Glu Thr Thr Ala Ser Thr

136 290 295 300

138 Ala Thr Ser Leu Val Met Phe Leu Gly Leu Asn Thr Glu Val Val Gln

139 305 310 315 320

141 Lys Val Arg Glu Glu Val Gln Glu Lys Val Glu Met Gly Met Tyr Thr

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/668,482

DATE: 08/30/2001

TIME: 15:04:31

Input Set : N:\Crf3\RULE60\09668482.txt

Output Set: N:\CRF3\08302001\I668482.raw

```

142          325          330          335
144 Pro Gly Lys Gly Leu Ser Met Glu Leu Leu Asp Gln Leu Lys Tyr Thr
145          340          345          350
147 Gly Cys Val Ile Lys Glu Thr Leu Arg Ile Asn Pro Pro Val Pro Gly
148          355          360          365
150 Gly Phe Arg Val Ala Leu Lys Thr Phe Glu Leu Asn Gly Tyr Gln Ile
151          370          375          380
153 Pro Lys Gly Trp Asn Val Ile Tyr Ser Ile Cys Asp Thr His Asp Val
154 385          390          395          400
156 Ala Asp Val Phe Pro Asn Lys Glu Glu Phe Gln Pro Glu Arg Phe Met
157          405          410          415
159 Ser Lys Gly Leu Glu Asp Gly Ser Arg Phe Asn Tyr Ile Pro Phe Gly
160          420          425          430
162 Gly Gly Ser Arg Met Cys Val Gly Lys Glu Phe Ala Lys Val Leu Leu
163          435          440          445
165 Lys Ile Phe Leu Val Glu Leu Thr Gln His Cys Asn Trp Ile Leu Ser
166          450          455          460
168 Asn Gly Pro Pro Thr Met Lys Thr Gly Pro Thr Ile Tyr Pro Val Asp
169 465          470          475          480
171 Asn Leu Pro Thr Lys Phe Thr Ser Tyr Val Arg Asn
172          485          490
175 (2) INFORMATION FOR SEQ ID NO: 3
176 (i) SEQUENCE CHARACTERISTICS:
177 (A) LENGTH: 1850 base pairs
178 (B) TYPE: nucleic acid
179 (C) STRANDEDNESS: single
180 (D) TOPOLOGY: linear
182 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3
184 TGTCGCCGTT GCTGTCGGTT GCTGTCGGAC GCTGTCTCCT CTCCAGAAGC TTGTTTTTCG      60
186 TTTTGGCGAT CAGTTGCGCG CTTCAAC ATG GGG CTG TAC ACC CTT ATG GTC ACC      114
187 Met Gly Leu Tyr Thr Leu Met Val Thr
188 1 5
190 TTT CTC TGC ACC ATC GTG CTA CCC GTT TTA CTC TTT CTC GCC GCG GTG      162
191 Phe Leu Cys Thr Ile Val Leu Pro Val Leu Leu Phe Leu Ala Ala Val
192 10 15 20 25
194 AAG TTG TGG GAG ATG TTA ATG ATC CGA CGA GTC GAT CCG AAC TGC AGA      210
195 Lys Leu Trp Glu Met Leu Met Ile Arg Arg Val Asp Pro Asn Cys Arg
196 30 35 40
198 AGT CCT CTA CCG CCA GGT ACC ATG GGC TTG CCG TTC ATT GGA GAA ACG      258
199 Ser Pro Leu Pro Pro Gly Thr Met Gly Leu Pro Phe Ile Gly Glu Thr
200 45 50 55
202 CTC CAG CTG ATC CTC CAG AGA AGG AAG TTT CTG CGC ATG AAA CGG CAG      306
203 Leu Gln Leu Ile Leu Gln Arg Arg Lys Phe Leu Arg Met Lys Arg Gln
204 60 65 70
206 AAA TAC GGG TGC ATC TAC AAG ACG CAC CTC TTC GGG AAC CCG ACT GTC      354
207 Lys Tyr Gly Cys Ile Tyr Lys Thr His Leu Phe Gly Asn Pro Thr Val
208 75 80 85
210 AGG GTG ATG GGA GCT GAT AAT GTG AGG CAG ATT CTG CTG GGC GAA CAC      402
211 Arg Val Met Gly Ala Asp Asn Val Arg Gln Ile Leu Leu Gly Glu His

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/668,482

DATE: 08/30/2001

TIME: 15:04:31

Input Set : N:\Crf3\RULE60\09668482.txt

Output Set: N:\CRF3\08302001\I668482.raw

212	90				95				100				105				
214	AAG	CTG	GTG	TCT	GTT	CAG	TGG	CCA	GCA	TCA	GTG	AGA	ACC	ATC	CTG	GGC	450
215	Lys	Leu	Val	Ser	Val	Gln	Trp	Pro	Ala	Ser	Val	Arg	Thr	Ile	Leu	Gly	
216					110				115				120				
218	TCT	GAC	ACC	CTC	TCC	AAT	GTC	CAT	GGA	GTT	CAA	CAC	AAA	AAC	AAG	AAA	498
219	Ser	Asp	Thr	Leu	Ser	Asn	Val	His	Gly	Val	Gln	His	Lys	Asn	Lys	Lys	
220				125					130				135				
222	AAG	GCC	ATT	ATG	AGG	GCG	TTC	TCT	CGA	GAT	GCT	CTG	GAG	CAC	TAC	ATT	546
223	Lys	Ala	Ile	Met	Arg	Ala	Phe	Ser	Arg	Asp	Ala	Leu	Glu	His	Tyr	Ile	
224			140					145				150					
226	CCC	GTG	ATC	CAG	CAG	GAG	GTG	AAG	AGC	GCC	ATA	CAG	GAA	TGG	CTG	CAA	594
227	Pro	Val	Ile	Gln	Gln	Glu	Val	Lys	Ser	Ala	Ile	Gln	Glu	Trp	Leu	Gln	
228		155				160					165						
230	AAA	GAC	TCC	TGC	GTG	CTG	GTT	TAT	CCA	GAA	ATG	AAG	AAA	CTC	ATG	TTT	642
231	Lys	Asp	Ser	Cys	Val	Leu	Val	Tyr	Pro	Glu	Met	Lys	Lys	Leu	Met	Phe	
232	170				175				180				185				
234	CGG	ATA	GCT	ATG	AGA	ATC	CTG	CTT	GGT	TTT	GAA	CCA	GAG	CAA	ATA	AAG	690
235	Arg	Ile	Ala	Met	Arg	Ile	Leu	Leu	Gly	Phe	Glu	Pro	Glu	Gln	Ile	Lys	
236				190					195				200				
238	ACG	GAC	GAG	CAA	GAA	CTG	GTG	GAA	GCT	TTT	GAG	GAA	ATG	ATC	AAA	AAC	738
239	Thr	Asp	Glu	Gln	Glu	Leu	Val	Glu	Ala	Phe	Glu	Glu	Met	Ile	Lys	Asn	
240			205					210				215					
242	TTG	TTC	TCC	TTG	CCA	ATC	GAC	GTT	CCT	TTC	AGT	GGT	CTG	TAC	AGG	GGT	786
243	Leu	Phe	Ser	Leu	Pro	Ile	Asp	Val	Pro	Phe	Ser	Gly	Leu	Tyr	Arg	Gly	
244			220				225					230					
246	TTG	AGG	GCA	CGC	AAT	TTC	ATT	CAC	TCC	AAA	ATT	GAG	GAA	AAC	ATC	AGG	834
247	Leu	Arg	Ala	Arg	Asn	Phe	Ile	His	Ser	Lys	Ile	Glu	Glu	Asn	Ile	Arg	
248		235				240					245						
250	AAG	AAA	ATT	CAA	GAT	GAC	GAC	AAT	GAA	AAC	GAA	CAG	AAA	TAC	AAA	GAC	882
251	Lys	Lys	Ile	Gln	Asp	Asp	Asp	Asn	Glu	Asn	Glu	gln	Lys	Tyr	Lys	Asp	
252	250				255				260				265				
254	GCC	CTT	CAG	CTG	TTG	ATC	GAG	AAC	AGC	AGA	AGA	AGT	GAC	GAA	CCT	TTT	930
255	Ala	Leu	Gln	Leu	Leu	Ile	Glu	Asn	Ser	Arg	Arg	Ser	Asp	Glu	Pro	Phe	
256			270					275				280					
258	AGT	TTG	CAG	GCG	ATG	AAA	GAA	GCA	GCT	ACA	GAG	CTT	CTA	TTT	GGA	GGT	978
259	Ser	Leu	Gln	Ala	Met	Lys	Glu	Ala	Ala	Thr	Glu	Leu	Leu	Phe	Gly	Gly	
260			285					290				295					
262	CAT	GAA	ACC	ACC	GCC	AGC	ACT	GCA	ACC	TCA	CTT	GTC	ATG	TTT	CTG	GGT	1026
263	His	Glu	Thr	Thr	Ala	Ser	Thr	Ala	Thr	Ser	Leu	Val	Met	Phe	Leu	Gly	
264			300				305					310					
266	CTG	AAC	ACA	GAA	GTG	GTG	CAG	AAG	GTC	AGA	GAG	GAG	GTT	CAG	GAG	AAG	1074
267	Leu	Asn	Thr	Glu	Val	Val	Gln	Lys	Val	Arg	Glu	Glu	Val	Gln	Glu	Lys	
268		315				320					325						
270	GTT	GAA	ATG	GGC	ATG	TAT	ACA	CCT	GGA	AAG	GGC	TTG	AGT	ATG	GAG	CTG	1122
271	Val	Glu	Met	Gly	Met	Tyr	Thr	Pro	Gly	Lys	Gly	Leu	Ser	Met	Glu	Leu	
272	330				335				340			345					
274	TTG	GAC	CAG	CTG	AAG	TAC	ACT	GGA	TGT	GTG	ATT	AAA	GAG	ACT	CTT	AGA	1170
275	Leu	Asp	Gln	Leu	Lys	Tyr	Thr	Gly	Cys	Val	Ile	Lys	Glu	Thr	Leu	Arg	
276				350				355				360					

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/668,482

DATE: 08/30/2001

TIME: 15:04:31

Input Set : N:\CrF3\RULE60\09668482.txt

Output Set: N:\CRF3\08302001\I668482.raw

```

278 ATC AAC CCT CCT GTT CCC GGA GGA TTC AGA GTC GCA CTC AAA ACC TTT      1218
279 Ile Asn Pro Pro Val Pro Gly Gly Phe Arg Val Ala Leu Lys Thr Phe
280           365           370           375
282 GAA TTG AAT GGT TAC CAA ATT CCT AAA GGA TGG AAC GTC ATT TAC AGC      1266
283 Glu Leu Asn Gly Tyr Gln Ile Pro Lys Gly Trp Asn Val Ile Tyr Ser
284           380           385           390
286 ATC TGT GAC ACG CAC GAT GTG GCC GAC GTC TTT CCA AAC AAA GAG GAG      1314
287 Ile Cys Asp Thr His Asp Val Ala Asp Val Phe Pro Asn Lys Glu Glu
288           395           400           405
290 TTC CAG CCG GAG AGA TTC ATG AGC AAA GGT CTG GAG GAC GGG TCC AGG      1362
291 Phe Gln Pro Glu Arg Phe Met Ser Lys Gly Leu Glu Asp Gly Ser Arg
292 410           415           420           425
294 TTT AAC TAC ATC CCC TTC GGA GGA GGA TCC AGG ATG TGT GTG GGC AAA      1410
295 Phe Asn Tyr Ile Pro Phe Gly Gly Gly Ser Arg Met Cys Val Gly Lys
296           430           435           440
298 GAG TTC GCC AAA GTG TTA CTC AAG ATC TTT TTA GTT GAG TTA ACG CAG      1458
299 Glu Phe Ala Lys Val Leu Leu Lys Ile Phe Leu Val Glu Leu Thr Gln
300           445           450           455
302 CAT TGC AAT TGG ATT CTC TCA AAC GGA CCC CCG ACA ATG AAA ACA GGC      1506
303 His Cys Asn Trp Ile Leu Ser Asn Gly Pro Pro Thr Met Lys Thr Gly
304           460           465           470
306 CCG ACT ATT TAC CCA GTG GAC AAT CTC CCT ACC AAA TTC ACT AGT TAT      1554
307 Pro Thr Ile Tyr Pro Val Asp Asn Leu Pro Thr Lys Phe Thr Ser Tyr
308           475           480           485
310 GTC AGA AAT TAGCCTAACC GGAGCTTTGT ACATATGTTT TTATTTTAGA      1603
311 Val Arg Asn
312 490
314 TGAAGTGTGA TGTATTGGAT ATTTTCTATT TTGTTTATAT AAAGCAGATG TGTATATAAG      1663
316 TCTATGCGAG GAAGCGAAAA CGAGGGCACT ACTTTCTCAT GGATCACTGT AATGCTACAG      1723
318 AGTGTCTGTG ATGTATATTT ATAATGTAGT TGTGTTATAT AGCTTTTGTA CTGTATGCAA      1783
320 CTTATTTAAC TCGCTCTTTA TCTCATGGGT TTTATTTAAT AAAACATGTT CTTACAAAAA      1843
322 AAAAAAA      1850
325 (2) INFORMATION FOR SEQ ID NO: 4
326 (i) SEQUENCE CHARACTERISTICS:
327 (A) LENGTH: 497 amino acids
328 (B) TYPE: amino acid
329 (C) STRANDEDNESS: single
330 (D) TOPOLOGY: linear
332 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4
334 Met Gly Leu Pro Ala Leu Leu Ala Ser Ala Leu Cys Thr Phe Val Leu
335 1           5           10           15
337 Pro Leu Leu Leu Phe Leu Ala Ala Ile Lys Leu Trp Asp Leu Tyr Cys
338           20           25           30
340 Val Ser Gly Arg Asp Arg Ser Cys Ala Leu Pro Leu Pro Pro Gly Thr
341           35           40           45
343 Met Gly Phe Pro Phe Phe Gly Glu Thr Leu Gln Met Val Leu Gln Arg
344           50           55           60
346 Arg Lys Phe Leu Gln Met Lys Arg Arg Lys Tyr Gly Phe Ile Tyr Lys
347 65           70           75           80

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/668,482

DATE: 08/30/2001

TIME: 15:04:32

Input Set : N:\Crf3\RULE60\09668482.txt

Output Set: N:\CRF3\08302001\I668482.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]